

MENU | SEARCH | INDEX | DETAIL

1/1



JAPANESE PATENT OFFICE

010
U.S. 1989
30912142/00
12/20

PATENT ABSTRACTS OF JAPAN

BEST AVAILABLE COPY

(11)Publication number: 07175891
(43)Date of publication of application: 14.07.1995

(51)Int.Cl.

G06K 9/34

(21)Application number: 05320766

(71)Applicant:

NIPPON TELEGR & TELEPH CORP
<NTT>

(22)Date of filing: 21.12.1993

(72)Inventor:

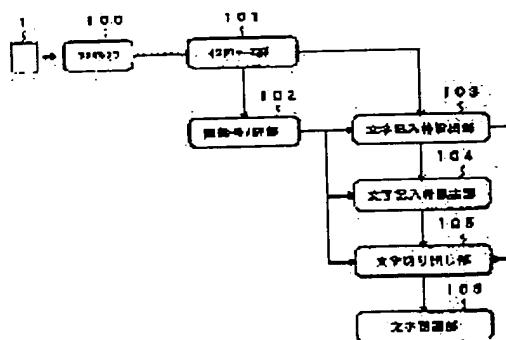
KOBAYASHI KIMITOMO
KITAMURA TADASHI
MIZUGAKI AKIO

(54) SLIP PROCESSOR

(57)Abstract:

PURPOSE: To easily use a slip printed by a printer or a copied slip by accurately recognizing the character description frame of the slip even when a described character is protruded from the character description frame or gets close to the character description frame in the case of sensing the character description frame together with the described character at the slip processor.

CONSTITUTION: An input slip 1 is read from a facsimile 100, and an image signal is stored in an image signal memory part 102. Based on a control mark, a character description frame detecting area larger than the character description frame is detected from this image signal by a character description frame detection part 103. Next, a fixed range from the upper, lower, right and left edges of that character description frame detecting area to the inside is investigated and only one line segment less than specified value width is erased by a character description frame erasure part 104 so that the phase shape of the described character can be kept without cutting or erasing the protruded part. Afterwards, the character is segmented by a character segmenting part 105 and as the result of this character description frame removal processing, the accurate character recognition is performed by a character recognizing part 106.



LEGAL STATUS

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平7-175891

(13) 公開日 平成7年(1995)7月14日

(51) Int.Cl.⁶

識別記号 庁内整理番号

F I

技術表示箇所

審査請求 未請求 請求項の数 1 OL (全 6 頁)

(21) 出願番号 特願平5-320766
(22) 出願日 平成5年(1993)12月21日

(71) 出願人 000004226
日本電信電話株式会社
東京都千代田区内幸町一丁目1番6号

(72) 発明者 小林 公知
東京都千代田区内幸町1丁目1番6号 日
本電信電話株式会社内

(72) 発明者 北村 正
東京都千代田区内幸町1丁目1番6号 日
本電信電話株式会社内

(72) 発明者 水書 章雄
東京都千代田区内幸町1丁目1番6号 日
本電信電話株式会社内

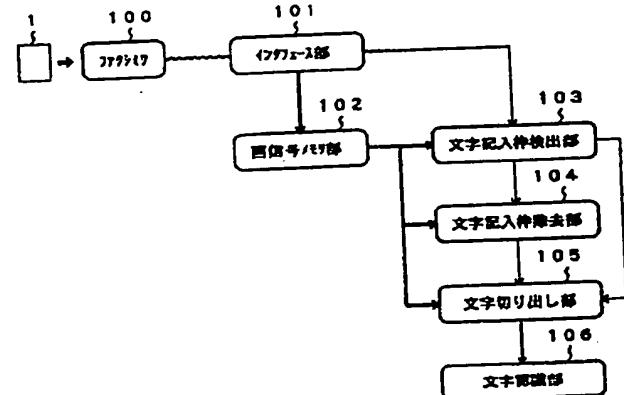
(74) 代理人 弁理士 志賀 富士弥

(54) 【発明の名称】 帳票処理装置

(57) 【要約】

【目的】 帳票処理装置にて帳票の文字記入枠が記入文字と共に感知される場合、記入文字が文字記入枠からみ出したり、文字記入枠に接近したりしても精度よく認識し、プリンタで印刷した帳票やコピーした帳票の使用を容易にする。

【構成】 ファクシミリ 100 から入力帳票 1 を読み取り、画信号を画信号メモリ部 102 へ格納する。この画信号から制御マニクをもとに文字記入枠より大きい文字記入枠検出領域を文字記入枠検出部 103 で検出する。次に、文字記入枠除去部 104 で、その文字記入枠検出領域の縁の上下左右から内側に一定の範囲を調べ、規定領域幅以下の 1 線分のみを除去することで、はみ出した部分が切断されたり消されたりせずに、記入文字の位相形状を保てるようにする。その後、文字切り出し部 105 で文字切り出しを行い、文字認識部 106 で上記の文字記入枠除去処理の結果として精度の良い文字認識を行う。



【特許請求の範囲】

【請求項 1】 スキャナで入力した入力帳票の文字情報を認識する文字認識装置において、文字記入欄を示す文字記入枠と該文字記入枠の副走査方向の位置を示す制御マークを前記スキャナに感知出来る色で記載した入力帳票を用い、前記入力帳票を前記スキャナに入力して得られる画信号中の前記制御マークの位置から前記文字記入枠より大きい文字記入枠検出領域を決定する第1の手段と、該第1の手段で決定した文字記入枠検出領域の縁の上下左右から内側へ一定の距離の範囲にある規定幅以下の1線分を除去する第2の手段と、該第2の手段で処理した文字記入枠検出領域中の文字を切り出して認識する第3の手段と、を有することを特徴とする帳票処理装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は、ファクシミリ等のスキャナで入力した入力帳票の文字情報を認識する帳票処理装置に関し、詳しくは入力帳票に記載した文字情報を正確に切り出すことのできる帳票処理装置に関するものである。

【0002】

【従来の技術】 一般に、文字認識で用いる入力帳票は、文字位置を示す制御マークが黒、文字記入枠がドロップアウトカラーで正確にOCR用紙に印刷される。しかし、ドロップアウトカラーを用いた2色刷り帳票は単価が高い。このため、帳票コストを下げるため、普通紙を用いてプリンタで印刷した帳票、または印刷した帳票からコピーした帳票を用いるようになってきた。

【0003】

【発明が解決しようとする課題】 しかしながら、このようなプリンタで印刷した帳票またはコピーした帳票では、制御マークから文字記入枠の位置を正確に算出し、文字記入枠を除去して文字記入枠内の文字を切り出し、認識する必要がある。その文字記入枠の除去にあたっては文字記入枠の傾斜や線幅のバラツキを考慮して文字記入枠の周辺領域も含め除去し、文字認識に当たっては文字記入枠内の文字を切り出して認識するため、それらの帳票に文字記入を行う場合、文字記入枠からはみ出さないように書くだけでなく、文字記入枠に近い所にも書いてはならないことが要求される。このため、書くのに神経を使うという問題があった。また、はみ出して記入した文字および文字記入枠に近い文字の部分は消されるため、認識出来ないという問題があった。さらに、プリンタおよびコピー等を用いて帳票を作成すると、文字記入枠の大きさおよび位置に変動がある場合、文字記入枠の除去位置自体も正確に検出出来ないため、文字記入枠中に書かれた文字が正確に切り出せないという問題があった。

【0004】 本発明は、上記問題点を解決するためにな

されたものであり、その目的は、帳票の文字記入枠が記入文字とともにスキャナに感知される場合において、記入文字が文字記入枠からはみ出したり、文字記入枠に接近したりしても、精度よく文字認識できるようにして、普通紙にプリンタで印刷した帳票やコピーした帳票の使用を容易にした帳票処理装置を提供することにある。

【0005】

【課題を解決するための手段】 上記の目的を達成するため、本発明の帳票処理装置においては、スキャナで入力した入力帳票の文字情報を認識する文字認識装置において、文字記入欄を示す文字記入枠と該文字記入枠の副走査方向の位置を示す制御マークを前記スキャナに感知出来る色で記載した入力帳票を用い、前記入力帳票を前記スキャナに入力して得られる画信号中の前記制御マークの位置から前記文字記入枠より大きい文字記入枠検出領域を決定する第1の手段と、該第1の手段で決定した文字記入枠検出領域の縁の上下左右から内側へ一定の距離の範囲にある規定幅以下の1線分を除去する第2の手段と、該第2の手段で処理した文字記入枠検出領域中の文字を切り出して認識する第3の手段と、を有することを特徴としている。

【0006】

【作用】 本発明の帳票処理装置では、ファクシミリ等のスキャナから入力した入力帳票から制御マークをもとに算出した文字記入枠よりも大きい文字記入枠検出領域を検出し、その文字記入枠検出領域の縁の上下左右から内側に一定の範囲を調べ、規定幅以下の1線分のみを除去することで、文字記入枠からはみ出した線分を切断したり消去したりするのを回避し、出来るだけ記入文字の位相形状を保つことが出来るようにより、帳票への記入文字が文字記入枠からはみ出したり、文字記入枠に接近したりしても、精度よく文字認識が行えるようにして、正読率を高く保てるようとする。

【0007】

【実施例】 以下、本発明の実施例を、図面を用いて詳しく説明する。

【0008】 図1は本発明の一実施例で使用する入力帳票の一例を示す図であり、1は入力帳票、2は文字記入枠の左の副走査位置を示す制御マーク、3は文字記入枠の右の副走査位置を示す制御マーク、4は文字記入枠、Nは1行の文字数である。なお、制御マーク2、3と文字記入枠4は、黒またはファクシミリ等のセンサで感知出来る色で印刷されている。そして、制御マーク2、3は入力帳票の両端に印刷し、この間に文字記入枠4を均等に配置する。

【0009】 図2は文字記入枠検出領域の検出方法を示した図であり、10は文字記入枠検出領域、dは制御マーク2と3の主走査距離、hは制御マーク2と3の副走査距離、Pnは文字記入枠の上部の中心位置である。

【0010】 次に、図1、図2を用いて文字記入枠検出

領域10の検出方法を示す。入力帳票1がファクシミリ等のスキャナで入力されて画信号メモリに格納されると(図示省略)、まず、入力帳票1の両端にある制御マーク2と3の検出が行われる。制御マーク2と3の検出は、画信号の先頭から、各走査線の両端から一定の範囲にある黒画素列を調べ、黒画素列が一定の範囲にあり副走査方向に一定の範囲連続したとき、制御マーク2または3が有りとして検出出来る。そして、一対の制御マーク2と3が検出されると、制御マーク2の右上端と制御マーク3の左上端の位置から制御マーク2と3の主走査距離dと副走査距離hを算出する。これらの値をもとに、以下の式により各行のn番目の文字記入枠の上部中心Pn(Xn, Yn)を算出する。ただし、検出した制御マーク2の右上端位置は(Mx, My)とする。

【0011】

$X_n = d \times n / N + Mx$, $Y_n = h \times n / N + My$
このように算出されたPnから文字記入枠4より大きな文字記入枠検出領域10を決める。

【0012】図3から図6までは文字記入枠の除去方法を示した図であって、図3は文字記入枠除去領域を検出する方法を示した図、図4(a), (b)は文字記入枠の横線の除去方法を示した図、図5(a), (b)は文字記入枠の縦線の除去方法を示した図、図6(a), (b), (c), (d)は文字枠除去例示した図であり、11は文字記入枠除去領域、FTは文字記入枠の上端位置、FBは文字記入枠の下端位置、FLは文字記入枠の左端位置、FRは文字記入枠の右端位置、AHは文字記入枠の横線除去領域、AVは文字記入枠の縦線除去領域、○は白画素、●は除去しない黒画素、◆は除去される黒画素、LVは横線除去領域内の縦線分、LHは縦線除去領域内の横線分である。

【0013】次に、これらの図3から図6までを用いて文字記入枠除去方法を説明する。まず、文字記入枠検出領域10が決まつたら、図3に示すように主走査方向および副走査方向に黒画素のヒストグラムをとり、副走査方向の上下から規定の値以上の黒画素数が最初に検出された位置をFTとFBとする。また同様に、主走査方向の左右から規定値以上の黒画素が最初に検出された位置をFLとFRとする。このようにして求めた文字記入枠の上端位置FT、文字記入枠の下端位置FB、文字記入枠の左端位置FL、文字記入枠の右端位置FRから入力帳票の傾斜と文字記入枠の線幅を考慮して決めた一定距離内側に入った位置の四角形から文字記入枠検出領域10の大きさまでを文字記入枠除去領域11とする。

【0014】次に、図4(a), (b)に示すように文字記入枠除去領域11のうち横線除去領域AH内を副走査方向(縦方向)に調べ、最初に2画素(1画素の白抜けを許容するため)の白画素で挟まれた規定値以下の黒画素列を除去すると、黒画素◆が除去できる。なお、横線除去領域AH内の縦線分LVは次の処理で除去され

る。続いて、図5(a), (b)に示すように文字記入枠除去領域11のうち縦線除去領域AV内を主走査方向(横方向)に調べ、最初に2画素(1画素の白抜けを許容するため)の白画素で挟まれた規定値以下の黒画素列を除去すると、黒画素◆が除去出来る。なお、縦線除去領域AV内の横線分LHは上記横線分除去処理で除去されている。このように横線除去領域AHと縦線除去領域AV内の1線分を除去すると文字記入枠4または相当する線分が除去出来る。以上のようにして文字記入枠4を除去した文字パターン例は図6(a), (b)に示すように、文字記入枠4と文字線分が重複しないかぎり文字の位相形状を保存出来ることがわかる。なお、図6

(c), (d)は、文字記入枠4と文字線分が重畳した場合の文字の位相形状が保存できない文字パターン例を示している。

【0015】図7(a), (b)は文字切り出し方法を示した図であり、20は文字を外接四角形で囲んだ文字領域、21はN文字分の送信用バッファ、22はi番目の文字書き込み領域である。次に、この図7を用いて文字パターンの切り出しを説明する。

【0016】まず、図2および図3～図6で説明した方法で文字記入枠検出領域10内の文字記入枠4が除去されると、図7(a)に示すようになる。この文字記入枠検出領域10の外側から内側に四角形で囲んでいき、規定数の黒画素列と接触する四角形を文字領域20として検出する。このように検出した文字領域20を切り出し、N文字分の送信用バッファ21のi番目の文字書き込み領域22の中央に配置されるように書き込む。

【0017】図8は本発明の実施例を示すブロック図であり、100は入力帳票1を走査するファクシミリ、101はファクシミリ100の画信号を取り込むインターフェース部、102は画信号を格納する画信号メモリ部、103は画信号中の文字記入枠4を検出する文字記入枠検出部、104は画信号中の文字記入枠4を除去する文字記入枠除去部、105は画信号中から文字を切り出し以下の文字認識部に転送する文字切り出し部、106は切り出した文字の認識を行う文字認識部である。

【0018】次に、図8の動作を説明する。まず、入力帳票1をファクシミリ100に入力する。ファクシミリ100は入力帳票1を走査してファクシミリ信号をインターフェース部101へ送信する。インターフェース部101はファクシミリ信号より画信号を取り出し、画信号メモリ部102へ格納する。格納が終了するとインターフェース部101は文字記入枠検出部103へ格納完了を通知する。インターフェース部101から格納完了通知を受けた文字記入枠検出部103は、図2で示した方法で文字記入枠検出領域10を検出し、文字記入枠検出領域10から文字記入枠除去領域11を検出する。1行分の文字記入枠除去領域11を検出すると個々の文字記入枠除去領域11の位置を文字記入枠除去部104へ、また、

文字記入枠検出領域 10 の位置を文字切り出し部 105 へ転送し、文字記入枠除去部 104 へ検出完了を通知する。検出完了通知を受けた文字記入枠除去部 104 は受信した 1 行分の文字記入枠除去領域 11 の位置情報をもとに、図 3 で説明した方法で個々の文字記入枠 4 を除去する。1 行分の文字記入枠 4 の除去が終了すると、除去完了通知を文字切り出し部 105 へ通知する。文字切り出し部 105 では、文字枠除去部 104 からの完了通知を受信すると、文字記入枠検出部 104 から受信した文字記入枠検出領域 10 の位置情報をもとに図 7 で示した方法で文字切り出しを行い、文字認識装置部 106 へ送信用バッファ 21 のデータを送信する。文字認識部 106 では、受信した送信用バッファ 21 のデータの中の 1 行分の文字を認識する。

【0019】以上の文字記入枠検出領域検出処理、文字記入枠除去処理、文字切り出し処理、文字認識処理を画信号の終わりまで行うことで、入力帳票 1 の処理が完了する。

【0020】なお、本実施例の文字記入枠検出領域検出方法、切り出し方法は一例であり、他の方法を用いてもよい。また、説明の処理パラメータも一例を示したものであり、処理対象画素数により異なる。

【0021】

【発明の効果】以上説明したように、本発明の帳票処理装置は、入力帳票の黒の文字記入枠または文字記入枠相当の線分を正確に除去して認識させるため、文字記入枠の近変を含めて除去する従来の方法と比べて処理も容易であり、かつてはみ出し文字の位相形状を保存するとともに文字記入枠に近い文字線分も消すことがないため入力帳票へ記入した文字の正読率を高く保てる。また、プリンタおよびコピー時のサイズ変化に対しても文字記入枠記載位置のバラツキを大幅に許容出来、正確な文字切り出しが出来るため、正読率を高く保てる。

【図面の簡単な説明】

【図 1】本発明の実施例で使用する入力帳票の一例を示した図

【図 2】上記実施例における文字記入枠検出領域の検出方法を示した図

【図 3】上記実施例の文字記入枠除去の方法における文字記入枠除去領域の検出方法を示した図

【図 4】(a), (b) は文字記入枠除去の方法における文字記入枠の横線の除去方法を示した図

【図 5】(a), (b) は文字記入枠除去の方法における文字記入枠の縦線の除去方法を示した図

【図 6】(a), (b), (c), (d) は文字記入枠除去の方法における文字記入枠除去結果の一例を示した図

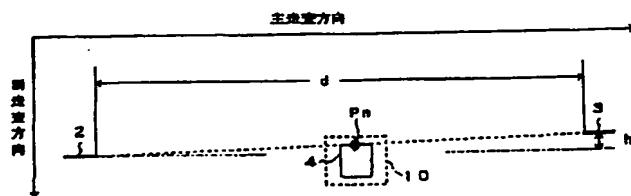
【図 7】(a), (b) は上記実施例における文字切り出し方法を示した図

【図 8】上記実施例の構成を示すブロック図

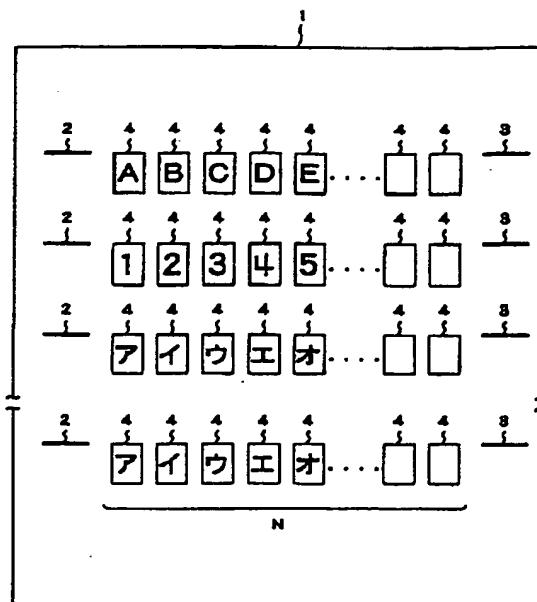
【符号の説明】

- 1 … 入力帳票
- 2 … 文字記入枠の左の副走査方向を示す制御マーク
- 3 … 文字記入枠の右の副走査方向を示す制御マーク
- 4 … 文字記入枠
- 10 … 文字記入枠検出領域
- 11 … 文字記入枠除去領域
- 20 … 文字パタンの外接四角形
- 21 … N 文字分の送信用バッファ
- 22 … i 文字目の文字書き込み領域
- 100 … ファクシミリ
- 101 … インタフェース部
- 102 … 画信号メモリ部
- 103 … 文字記入枠検出部
- 104 … 文字記入枠除去部
- 105 … 文字切り出し部
- 106 … 文字認識部
- P a … 文字記入枠の上部の中心位置
- P L … 文字記入枠の上端位置
- F R … 文字記入枠の下端位置
- F T … 文字記入枠の左端位置
- F B … 文字記入枠の右端位置
- A H … 文字枠の横線除去領域
- A V … 文字記入枠の縦線除去領域
- L V … 横線除去領域内の縦線分
- L H … 縦線除去領域内の横線分

【図 2】

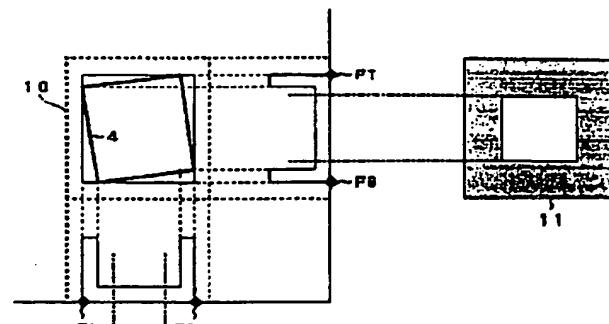


【図1】

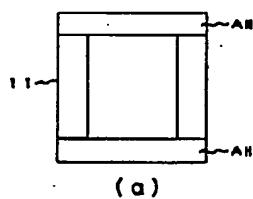


1 … 入力機票
 2 … 文字記入枠の左の前進検査方向を示す制御マーク
 3 … 文字記入枠の右の前進検査方向を示す制御マーク
 4 … 文字記入枠

【図3】

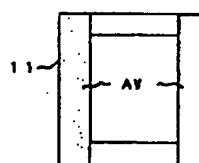


【図4】

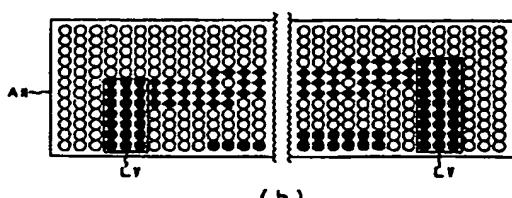


(a)

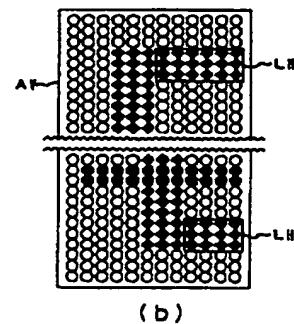
【図5】



(a.)

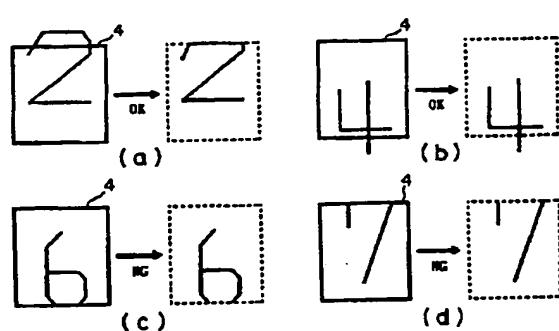


(b)

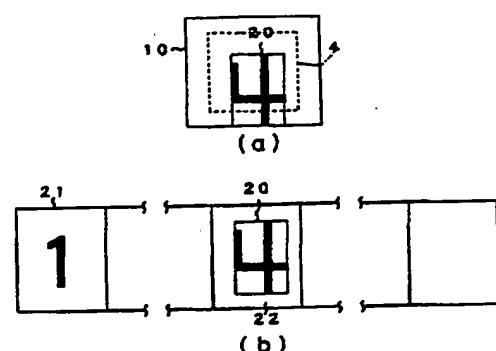


(b)

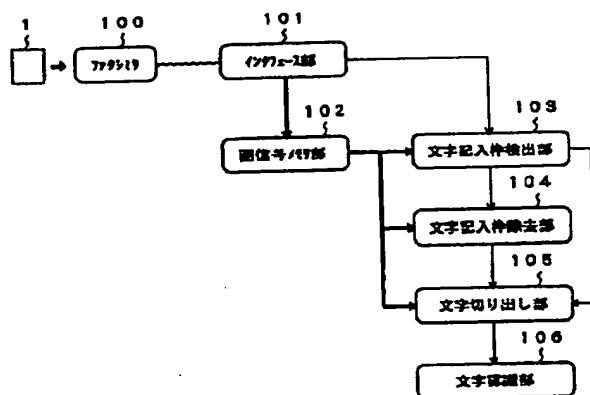
【図6】



【図7】



【図8】



**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.